

Analysis of Democratic Voting Principles used in Distributed Greedy Algorithms

Dennis Sundman, *Student member, IEEE*, Saikat Chatterjee, *Member, IEEE*,
and Mikael Skoglund, *Senior Member, IEEE*

Abstract—A key aspect for any greedy pursuit algorithm used in compressed sensing is a good support-set detection method. For distributed compressed sensing, we consider a setup where many sensors measure sparse signals that are correlated via the existence of a signals' intersection support-set. This intersection support-set is called the joint support-set. Estimation of the joint support-set has a high impact on the performance of a distributed greedy pursuit algorithm. This estimation can be achieved by exchanging local support-set estimates followed by a (consensus) voting method. In this paper we endeavor for a probabilistic analysis of two democratic voting principle that we call majority and consensus voting. In our analysis, we first model the input/output relation of a greedy algorithm (executed locally in a sensor) by a single parameter known as probability of miss. Based on this model, we analyze the voting principles and prove that the democratic voting principle has a merit to detect the joint support-set.

Index Terms—Greedy algorithms, distributed detection, hard decision.

I. INTRODUCTION

Compressed sensing (CS) [1], [2] typically considers a single-sensor scenario, where the main task is reconstruction of a large-dimensional signal-vector from a small-dimensional measurement-vector by using a-priori knowledge that the signal is sparse in a known domain. Several CS reconstruction algorithms have been considered in the literature, for example convex optimization- [3], [4], Bayesian- [5], [6] and greedy pursuit (GP) algorithms. The greedy pursuit (GP) algorithms are popular due to their low complexity and good performance. From a measurement vector, the GP algorithms use linear algebraic tools to estimate the underlying *support-set* of the sparse signal-vector followed by estimating associated signal values; here we mention that good support-set estimation is a key aspect for the GP algorithms. A few examples of typical GP algorithms include: matching pursuit [7], orthogonal matching pursuit (OMP) [8], CoSaMP [9], subspace pursuit (SP) [10], but there are many others [11], [12], [13], [14], [15]. For the GP algorithms, just as for any CS reconstruction algorithm, providing analytical performance guarantees is an important yet challenging task. These guarantees are typically done through worst case analysis based on restricted isometry property [16] and mutual coherence inequalities.

Distributed (or de-centralized) CS (DCS) refers to a problem of multiple sensors connected over a network, where the

sensors observe correlated sparse signals through CS measurements. By the term DCS we refer both to simultaneous estimation in a distributed network [17], [18], [19], [20], [21], [22] and to multiple measurement vector setups in some fusion center [3], [23], [24], [25]. Recently we developed several GP algorithms for DCS, called distributed greedy pursuits [26], [27], [28], [29]. In DCS, two (of many) models for signal correlations are the common support-set model and the mixed support-set model [30]. In the common support-set model, the same (joint) *full* support-set is assumed for all signals measured at different sensors, while in the mixed support-set model a joint *partial* support-set is assumed for all sensors. Based on these models, a prominent approach for distributed GP algorithms is to let the sensors in the network exchange (or transmit to a centralized point) full support-set estimates and then, using only support-set knowledge, estimate the joint support-set. A better estimate of the joint support-set can then be used to improve DCS reconstruction performance.

In general, theoretical performance analysis of distributed GP algorithms is non-trivial and we recently developed DIPP (distributed parallel pursuit) – a distributed greedy pursuit algorithm – with such theoretical guarantees in [29]. Through analysis and simulations we have shown that DIPP performs better than local GP algorithms, such as SP. In DIPP and other distributed greedy pursuits, the joint support-set is estimated by a consensus voting method. In several of our earlier works [27], [28], we assumed that democratic based voting is suitable for consensus, and in [29] we proved theoretical reconstruction guarantees based on this assumption. The advantage of voting has earlier been studied in politics and finance as early as 1785 [31], [32]. In this paper, we endeavor to prove that the assumption of democratic voting for support-set estimation, based on GP algorithms, indeed has a merit. In our approach, we assume that support-sets estimated from GP algorithms executed locally in several sensors likely contain independent errors. Therefore, based on probability of detection, miss, and false alarms, we first model the input/output relation of relevant GP algorithms by using standard detection theory framework.

Using the input/output relation, we provide probability results for consensus strategies based on democratic voting applied in scenarios that employ the common and mixed support-set models. The main contributions of this paper can be summarized as:

- Defining the input/output relation of relevant GP algorithms.
- Probabilistic analysis of democratic voting used in dis-

tributed GP algorithms for both common and mixed support-set models.

The outline of the paper is as follows: We first introduce some notation in the next subsection. Then, in Section II, we introduce the signal model, the common support-set, and mixed support-set models. In Section III, we introduce an input/output relation to model single sensor performance, which is then used for analyzing different voting strategies for the common support-set model in Section IV and the mixed support-set model in Section V. Then, in Section VI, we provide experimental verification of the results achieved.

A. Notation

Sets are denoted by calligraphic capital letters, in particular; \mathcal{T} , \mathcal{I} and \mathcal{J} are support-sets or partial support-sets. We define the full set $\Omega = \{1, 2, \dots, N\}$ and the set complement $\mathcal{J}^c = \Omega \setminus \mathcal{J}$, where ' \setminus ' is the set-minus. We denote the event of an index i residing in the support-set \mathcal{T} by $i \in \mathcal{T}$. If i resides in two support-sets \mathcal{T}_p and \mathcal{T}_q – we use either $(i \in \mathcal{T}_p, i \in \mathcal{T}_q)$ or $i \in (\mathcal{T}_p \cap \mathcal{T}_q)$; where the one providing most insight will be used. The probability of an index i residing inside the support-set \mathcal{T} is denoted by $\mathbb{P}(i \in \mathcal{T})$. Lastly, we denote the conditional probability, where by $\mathbb{P}(i \in \mathcal{J} | i \in \mathcal{T})$ we refer to the probability that an index i be in \mathcal{J} , given that i is (randomly) in \mathcal{T} . Lastly we introduce two algorithmic notations

$$\text{vote}_1(\mathbf{z}, \mathcal{T}) \triangleq \{\forall i \in \mathcal{T}, \text{perform } z_i = z_i + 1\}. \quad (1)$$

$$\text{max_indices}(\mathbf{z}, T) \triangleq \{\text{select the } T \text{ largest amplitude indices of } \mathbf{z}\}. \quad (2)$$

II. SYSTEM MODEL

In this section we define the distributed compressed sensing (DCS) problem, the common support-set model and the mixed support-set model.

A. Distributed Compressed Sensing

In distributed compressed sensing (DCS), each p 'th sensor measures a signal $\mathbf{x}_p \in \mathbb{R}^N$ through the following linear relation

$$\mathbf{y}_p = \mathbf{A}_p \mathbf{x}_p + \mathbf{e}_p, \quad \forall p \in \mathcal{L}, \quad (3)$$

where $\mathbf{y}_p \in \mathbb{R}^M$ is a measurement vector, $\mathbf{A}_p \in \mathbb{R}^{M \times N}$ is a measurement matrix, $\mathbf{e}_p \in \mathbb{R}^M$ is some measurement noise and \mathcal{L} is a global set containing all sensors (nodes) in the network ($|\mathcal{L}| = L$). The signal vector $\mathbf{x}_p = [x_p(1) \ x_p(2) \ \dots \ x_p(N)]$ is T -sparse, meaning it has T elements that are non-zero. Thus, the setup describes an under-determined system, where $T < M < N$. The element-indices corresponding to non-zero values are collected in the support-set \mathcal{T}_p ; that means $\mathcal{T}_p = \{i : x_p(i) \neq 0\}$ and $|\mathcal{T}_p| = T$. A dense vector containing only the non-zero values of \mathbf{x}_p is represented by $\mathbf{v}_p = [x_p(\mathcal{T}_p(1)), x_p(\mathcal{T}_p(2)), \dots, x_p(\mathcal{T}_p(T))]$, which may also be independent locally and across the network. Throughout this paper we use measurement matrices that have unit

ℓ_2 -norm columns and characterize the signal- to-noise ratio using the signal-to-measurement-noise-ratio (SMNR), which is defined for sensor p as

$$\text{SMNR} \triangleq \frac{\mathcal{E}\{\|\mathbf{x}_p\|_2^2\}}{\mathcal{E}\{\|\mathbf{e}_p\|_2^2\}}. \quad (4)$$

Furthermore, \mathbf{A}_p and \mathbf{e}_p are independent both locally and across the network.

In order to benefit from cooperation in the network, some correlation in the signal vector \mathbf{x}_p must be present. In the following two subsections we present these correlations by introducing the common and mixed support-set models.

B. Common Support-set Model

In the common support-set model [33], [30], the support-sets of all signals in the network \mathbf{x}_p are identical. That is

$$\mathcal{T}_p = \mathcal{J} \quad \forall p \in \mathcal{L}, \quad (5)$$

where we refer to \mathcal{J} as the *joint* support-set.

C. Mixed Support-set Model

A natural extension to the common support-set model is the mixed support-set model, proposed by us in [22], [27], [30]. In this case there exists an intersection between all support-sets \mathcal{T}_p . Denoting $\mathcal{J} = \cap_{p \in \mathcal{L}} \mathcal{T}_p$, we have

$$\mathcal{T}_p = \mathcal{I}_p \cup \mathcal{J} \quad \forall p \in \mathcal{L}. \quad (6)$$

Here, we refer to \mathcal{J} as the joint part of the support-set (or simply joint support-set) and $\mathcal{I}_p \triangleq \mathcal{T}_p \setminus \mathcal{J}$ is the individual part.

Assumption 1: Denoting $|\mathcal{T}_p| = I \ \forall p$ and $|\mathcal{J}| = J$, the following assumptions are used throughout the paper:

- 1) Elements of support-sets are uniformly distributed,

$$\mathbb{P}(i \in \mathcal{T}_p) = \frac{|\mathcal{T}_p|}{|\Omega|} = \frac{T}{N}, \quad \forall p \in \mathcal{L}. \quad (7)$$

- 2) $\mathcal{I}_p \cap \mathcal{J} = \emptyset, \quad \forall p \in \mathcal{L}.$
- 3) $\mathcal{I}_p \cap \mathcal{I}_q = \emptyset, \quad \forall p, q \in \mathcal{L}, p \neq q.$
- 4) Hence, $T = I + J.$ □

III. MODELING THE INPUT/OUTPUT RELATION FOR GREEDY PURSUITS

A GP algorithm in sensor p will attempt to find the true support-set \mathcal{T}_p . Influencing the chances of success are a number of factors: signal amplitudes (i.e., \mathbf{v}_p), measurement noise \mathbf{e}_p , sparsity T and measurement matrix \mathbf{A}_p realization. Illustrated in Fig. 1, is the whole procedure from an underlying \mathcal{T}_p , signal acquisition according to (3), to recovered support-set estimate $\hat{\mathcal{T}}_p$. In Fig. 2a, we have simplified the previous figure in one box, referred to as the *System*. Borrowing terms from detection theory we model the system (see Definition 1), where the idea is to replace the factors influencing support-set recovery performance with one single parameter ϵ_p , shown in Fig. 2b. Introduction of this single parameter helps to bring analytical tractability, which we will witness in Sections IV and V.

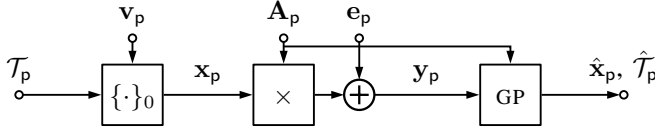


Fig. 1. The CS system considered in this paper. From known underlying support-set to estimate provided by reconstruction algorithm.

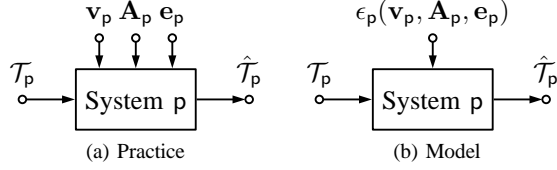


Fig. 2. Two simplified figures of the full system.

Definition 1 (System model): The support-set estimate \hat{T}_p of any unbiased GP algorithm described in Fig. 2b follows

$$\mathbb{P}(i \in \hat{T}_p) = \frac{T}{N} \quad (8)$$

$$\mathbb{P}(i \in \hat{T}_p | i \in T_p) = \mathbb{P}(\text{"detect"}) = 1 - \epsilon_p \quad (9)$$

$$\mathbb{P}(i \notin \hat{T}_p | i \in T_p) = \mathbb{P}(\text{"miss"}) = \epsilon_p \quad (10)$$

$$\mathbb{P}(i \in \hat{T}_p | i \in T_p^c) = \mathbb{P}(\text{"false alarm"}) = \frac{T}{N-T} \epsilon_p, \quad (11)$$

where $0 \leq \epsilon_p \leq \frac{N-T}{N}$. Observe that $i \notin \hat{T}_p = i \in \hat{T}_p^c$. These probabilities should be read as, for example in (9): “The probability that index i is part of the output \hat{T}_p from the system, provided that this index is already part of the true underlying support-set T_p ”. For the remainder of the paper, we assume that all sensors in the network have statistically identical system and signal conditions, meaning that $\epsilon_p = \epsilon \forall p$. \square

Discussion: The input/output relation in Definition 1 follows from symmetry arguments. Since the system is symmetric and T_p is uniformly random, any unbiased (fair) reconstruction algorithm will produce \hat{T}_p which is also uniformly random (8); an unbiased algorithm should not favor any correct index over than any other correct index, resulting in (9). Similarly, the algorithm should not favor any missed index over another missed index (10). Furthermore, whenever a support-index is missed, a false alarm has occurred; therefore the false alarm can be parametrized by the same parameter as the probability of miss and detect (11). Using ϵ to specify the system behavior, we see that the worst possible system will select indices for the support-set uniformly at random. Thus the upper-bound on ϵ is $\epsilon_{\max} = \frac{N-T}{N}$, which means that the worst $\mathbb{P}(\text{"detect"}) = \frac{T}{N}$. \square

At this point, there is no closed-form expression of the parameter ϵ as it would require complete characterization of A_p , x_p , e and of the present GP algorithm. Such an analysis is outside the scope of this paper; instead we estimate ϵ through experiments. This can in practice be performed, for example, by using pilot signals. We now present the first result.

Proposition 1: The probability that an index ‘ i ’ is correct for sensor p provided that it is found by the p ’th system is given by

$$\mathbb{P}(i \in T_p | i \in \hat{T}_p) = \mathbb{P}(i \in \hat{T}_p | i \in T_p). \quad (12)$$

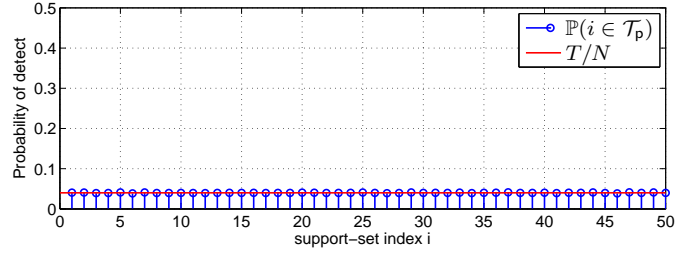


Fig. 3. Figure showing how often each index occurs in the output of the system, based on (8) of the system model.

Proof:

$$\begin{aligned} \mathbb{P}(i \in T_p | i \in \hat{T}_p) &\stackrel{(a)}{=} \frac{\mathbb{P}(i \in \hat{T}_p | i \in T_p) \mathbb{P}(i \in T_p)}{\mathbb{P}(i \in \hat{T}_p)} \\ &\stackrel{(b)}{=} \frac{\mathbb{P}(i \in \hat{T}_p | i \in T_p) \frac{T}{N}}{\frac{T}{N}} \\ &= \mathbb{P}(i \in \hat{T}_p | i \in T_p), \end{aligned}$$

where we in (a) have used Bayes’ rule and in (b) have used (7) and (8). \blacksquare

A. Numerical Verification of the System Model

In order to verify the system model in Definition 1, we perform two different tests. As GP algorithm we have used the well known subspace pursuit (SP) algorithm, however; similar results can be obtained with other GP algorithms that are based on fixed support-set size.

In the first test, presented in Fig. 3, we verify (8). The test is based on 10^5 random: support-sets T_p , signal realizations v_p , measurement matrices A_p and noises e_p (generated such that SMNR = 20 dB). In Fig. 3, $N = 50$ and $T = 2$ to make the outcome observable ($M = 7$). Along the x-axis we show the support-set index i , and on the y-axis, we show how many times each index appears in the output \hat{T}_p . From this figure, we see that by using the proposed setup, the output from the algorithm is uniform, which verifies (8) of the definition.

In the second test, presented in Fig. 4; (9), (10) and (11) are verified where $N = 50$ and $T = 2$ (and $M = 7$). Here, there are 10^5 random: signal realizations v_p , measurement matrices A_p and noises e_p (such that SMNR = 20 dB). The support-set $T_p = [14, 26]$ is fixed in order to produce an informative figure. Along the x-axis is the support-set index i , and on the y-axis, we show how many times each index appears in the output \hat{T}_p . We can now directly identify the three equations (9), (10) and (11). First, we estimate ϵ by counting the number of false alarms (11); in this case $\hat{\epsilon} = 0.267$. Then (9) and (10) are found directly from $\hat{\epsilon}$. We will now apply the input/output relation model to more complex scenarios.

IV. VOTING BASED DETECTION FOR THE COMMON SUPPORT-SET MODEL

In this section we introduce the concept of voting based on support-set estimates from a number of nodes. We model signal correlation according to the common support-set model

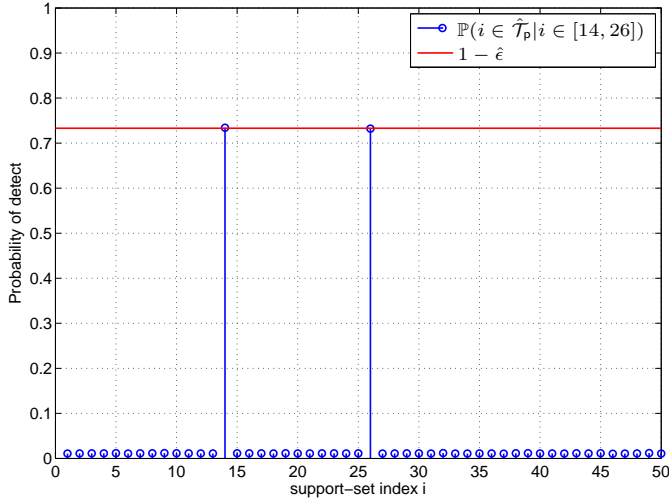


Fig. 4. Simulation verification that indeed there exist an underlying ϵ such that the proposed system model holds.

(see Section II-B). Throughout this section we use \mathcal{T} and \mathcal{J} interchangeably for the same purpose, since they are equivalent in the common support-set model.

Consider a setup with multiple sensor nodes where each sensor gathers CS measurements and runs a local GP algorithm to find a local support-set estimate. The support-set estimates from all nodes are then sent to a fusion center (or exchanged distributively) for estimation of \mathcal{J} .

A. Algorithm

We propose a fusion center strategy based on democratic voting where, assuming T is known, the strategy for the final estimate is to choose the T indices with most votes. This is a majority voting strategy and a detailed description is presented in Algorithm 1.

Algorithm 1 majority: Executed in the local node p

Input: $\{\hat{\mathcal{T}}_p\}_{p \in \mathcal{L}}, T$

Initialization: $\mathbf{z} \leftarrow \mathbf{0}_{N \times 1}$

Algorithm:

- 1: **for** each $p \in \mathcal{L}$ **do**
- 2: $\mathbf{z} \leftarrow \text{vote}_1(\mathbf{z}, \hat{\mathcal{T}}_p)$ (The estimate of sensor p)
- 3: **end for**
- 4: $\hat{\mathcal{J}} \leftarrow \text{max_indices}(\mathbf{z}, J)$

Output: $\hat{\mathcal{J}}$ (observe that this is an estimate of $\mathcal{T} = \mathcal{J}$)

Studying Algorithm 1, we see that the inputs are the support-set estimates from all sensors in the network, and the support-set cardinality. In the initialization phase, a large N -sized vector \mathbf{z} is created; where the votes of the sensors are collected. Then, the estimate $\hat{\mathcal{J}}$ is chosen based on the highest T values in \mathbf{z} , which corresponds to majority voting. Observe that when knowledge is available about J , the majority may be used also for the mixed support-set model, which we did (under another name) in [27].

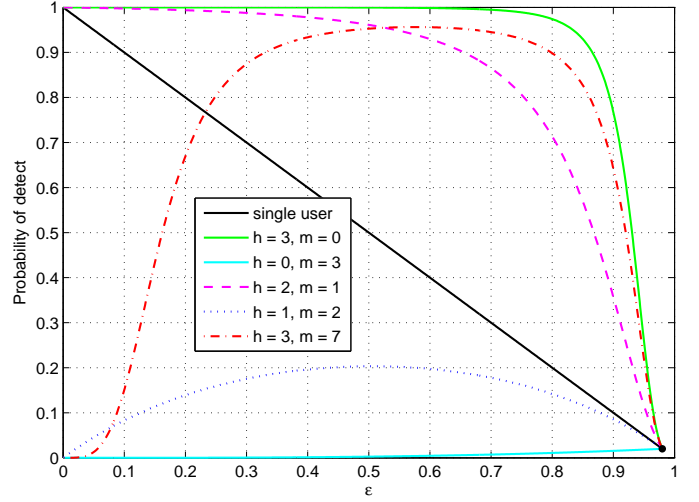


Fig. 5. A few examples of probability of detect for the majority algorithm using the common support-set model.

B. Analysis

When the nodes have found the support-set estimates by GP algorithms, we use the input/output relation in Definition 1 to provide some fundamental results valid for the majority algorithm.

Proposition 2: In a setup with $L = h + m$ sensors with signal support-sets $\hat{\mathcal{T}}_{p_l}$ for $l = 1, 2, \dots, h$ and $\hat{\mathcal{T}}_{q_l}$ for $l = 1, 2, \dots, m$, let us assume that the index $i \in \left(\bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c\right)$. Then, the majority algorithm finds the estimate $\hat{\mathcal{J}}$ such that $i \in \hat{\mathcal{J}}$. In this case, the probability of detection is

$$\begin{aligned} \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ = \frac{(1 - \epsilon)^h \epsilon^m J}{(1 - \epsilon)^h \epsilon^m J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N - J)}, \end{aligned} \quad (13)$$

where $J = T$.

Proof: Proof in Appendix A-A. ■

Getting any insight for the behavior of majority from Proposition 2 is a non-trivial since (13) is a complicated function of m, h, J, T and N . For better understanding, we provide an example where some parameters are fixed.

Example 1: Using $N = 1000, T = J = 20$, we provide Fig. 5 where several pairs of $\{h, m\}$ are tested via (13). The black curve corresponds to the disconnected performance of Proposition (1) and the black dot corresponds to the probability of detect at $\epsilon_{\max} = \frac{N-T}{N}$ which is the biggest value ϵ can take. Worth noticing in this figure is the interplay between hits and misses which may cause the performance to be very good at some parts, while being poor at other parts. This is illustrated in the curve for $h = 3, m = 7$. An observation we found is that whenever $h > m$ we get good performance. □

Using majority voting, it is intuitively clear that more votes are always better (for a constant number of total sensors in the network). We show this explicitly with the following proposition.

Proposition 3: For the same setup as in Proposition 2 the following relation holds

$$\begin{aligned} \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ \geq \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^{h-1} \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{q_l}^c) \end{aligned} \quad (14)$$

Proof: Proof in Appendix A-A. ■

By proposition 3, it is clear that in a network of sensors, under the common support-set model, the majority voting has a merit to detect the support-set.

V. MIXED SUPPORT-SET: DISTRIBUTED PARALLEL PURSUIT

We now consider the voting approach in a scenario where the signal correlation is modeled with the mixed support-set model. Assuming T to be known (but not J), we previously developed such an algorithm in [29], where it is called consensus voting. In this case, there is no fusion center; instead the sensors exchange support-set estimates and apply the consensus algorithm locally, based on information from the neighboring sensors.

A. Algorithm

The consensus algorithm differs from majority since it has no knowledge of the support-set size of the joint support J . Instead it performs a threshold operation by selecting components with at least two votes. We have provided consensus in Algorithm 2.

Algorithm 2 consensus: Executed in the local node p

Input: $\{\hat{\mathcal{T}}_q\}_{q \in \mathcal{L}_p^{\text{in}}}, \hat{\mathcal{T}}_p, T$

Initialization: $\mathbf{z} \leftarrow \mathbf{0}_{N \times 1}$

Algorithm:

- 1: $\mathbf{z} \leftarrow \text{vote}_1(\mathbf{z}, \hat{\mathcal{T}}_p)$ (p-th node's estimate)
- 2: **for** each $q \in \mathcal{L}_p^{\text{in}}$ **do**
- 3: $\mathbf{z} \leftarrow \text{vote}_1(\mathbf{z}, \hat{\mathcal{T}}_q)$ (The neighbors' estimates)
- 4: **end for**
- 5: Choose $\hat{\mathcal{J}}_p$ such that $(z(i) \geq 2) \forall i \in \hat{\mathcal{J}}_p$ and $|\hat{\mathcal{J}}_p| \leq T$

Output: $\hat{\mathcal{J}}_p$

Studying algorithm 2, the inputs are: a set of estimated support-sets $\{\hat{\mathcal{T}}_q\}_{q \in \mathcal{L}_p^{\text{in}}}$ from the neighbors, the locally estimated support-set $\hat{\mathcal{T}}_p$, and the sparsity level T . The estimate of $\hat{\mathcal{J}}_p$ is formed (step 5) such that no index in $\hat{\mathcal{J}}_p$ has less than two votes (i.e., each index in $\hat{\mathcal{J}}_p$ is present in at least two support-sets from $\{\{\hat{\mathcal{T}}_q\}_{q \in \mathcal{L}_p^{\text{in}}}, \hat{\mathcal{T}}_p\}$)¹. If the number of indices with at least two votes exceed the cardinality T , we pick the T largest indices. Thus, the consensus strategy can be summarized as:

¹For node p, this is equivalent to let algorithm choose $\hat{\mathcal{J}}_p$ as the union of all pair-wise intersections of support-sets (see the analysis section V-B for details).

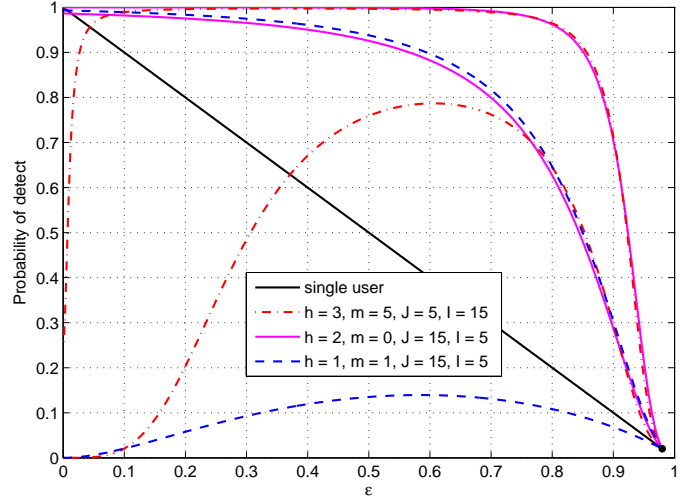


Fig. 6. Analytical results for the mixed support-set model. Observe here that there is always a total number of $h + m + 1$ nodes present for a consensus algorithm.

- Pick indices for $\hat{\mathcal{J}}_p$ that have two votes
- If $|\hat{\mathcal{J}}_p| > T$, choose the T largest indices

In the following we will analyze the consensus strategy using the input/output relation of Definition 1.

B. Analysis

Assuming the nodes use GP algorithms to find the support-set estimates, we obtain the following results.

Proposition 4: The probability that an index ' i ' is correct for sensor ' p ', provided that this index is detected by the sensor ' p ' itself and additionally ' h ' neighbors, but not detected by ' m ' neighbors is given by (15).

Proof: Proof in Appendix A-B. ■

Proposition 5: The probability that an index ' i ' is correct for sensor ' p ', provided that this index is detected by ' h ' neighbors, but not detected by the sensor ' p ' itself and additionally ' m ' neighbors is given by (16).

Proof: Proof in Appendix A-B. ■

Getting any insight from these propositions is difficult. Therefore, we provide the following numerical example.

Example 2: In Fig. 6 we provide examples for the mixed support-set model using Proposition 4 and Proposition 5. In this system $N = 1000$ and ϵ is varied. Notice in Fig. 6, that there are two curves for each configuration. The top-most curve corresponds to (15), where the sensor ' p ' itself found the index, and the lower-most curve corresponds to (16), where the sensor ' p ' itself missed the index. By testing it seems that, similarly to majority voting, when $h > m$, the performance is good.

Derivation of further general results based on Proposition 4 and Proposition 5, for example providing general precise requirements under which consensus provides higher probability than a single sensor case is non-trivial. Instead, we assume a limited number of neighbors and fix a number of parameters according to [29]. In particular, we assume that each local node p has two independent neighbors. Using

$$\mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) = \frac{(1-\epsilon)^{h+1} \epsilon^m \frac{J}{N} + (1-\epsilon) \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{I}{N}}{(1-\epsilon)^{h+1} \epsilon^m \frac{J}{N} + (h+1)(1-\epsilon) \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{I}{N} + m\epsilon \left(\frac{T}{N-T}\epsilon\right)^{h+1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m-1} \frac{I}{N} + \left(\frac{T}{N-T}\epsilon\right)^{h+1} \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{N-J-(m+h+1)I}{N}}. \quad (15)$$

$$\mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) = \frac{(1-\epsilon)^h \epsilon^{m+1} \frac{J}{N} + \epsilon \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{I}{N}}{(1-\epsilon)^h \epsilon^{m+1} \frac{J}{N} + h(1-\epsilon) \left(\frac{T}{N-T}\epsilon\right)^{h-1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} \frac{I}{N} + (m+1)\epsilon \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{I}{N} + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} \frac{N-J-(m+h+1)I}{N}}. \quad (16)$$

information obtained by the neighbors and the local node, the consensus endeavors to estimate the joint support part as $\hat{\mathcal{J}}_p$. Following algorithm 2, we note that in this case

$$\hat{\mathcal{J}}_p = \left\{ i : i \in \left((\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r) \cup (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r^c) \cup (\hat{\mathcal{T}}_p^c \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r) \right) \forall q, r \in \mathcal{L}_p^{\text{in}}, q \neq r \right\}. \quad (17)$$

Using (17), we provide the following remark in numerical manner.

Remark 1: When $N = 1000$, $T = 20$, $J = 15$, $V = 2$ and $0.0140 \leq \epsilon \leq \epsilon_{\max} = 0.98$, then the probability of $i \in \hat{\mathcal{J}}_p$ to be correct is always bigger than or equal to the probability of $i \in \hat{\mathcal{T}}_p$ to be correct, that is

$$\mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{J}}_p) \geq \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p). \quad (18)$$

Proof: Proof in Appendix A-B. ■

We note that although Remark 1 strongly suggest that the majority voting provides for a good result, we can consider the typical CS condition that N is very large. Then an even stronger result can be formulated as in the following corollary.

Corollary 1: If $J \geq 1$, T grows sublinearly in N , and $\hat{\mathcal{J}}_p$ is the output of consensus, then

$$\lim_{N \rightarrow \infty} \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{J}}_p) = 1. \quad (19)$$

Proof of Corollary 1: For this proof, we show that (15) tends to one when $h \geq 1$ (follows from $\hat{\mathcal{J}}_p$) and that (16) tends to one when $h \geq 2$ (also follows from $\hat{\mathcal{J}}_p$).

First consider (15) and note that it can never happen that $m > 0$ when $\epsilon = 0$. Then it is straight-forward to see that $\left(\frac{T}{N-T}\epsilon\right)^h \rightarrow 0$ since $h \geq 1$, hence the whole expression tends to 1.

Similarly for (16), note that it can never happen that $m > 0$ when $\epsilon = 0$. Then it is straight-forward to see that $\left(\frac{T}{N-T}\epsilon\right)^{h-1} \rightarrow 0$ since $h \geq 2$, consequently the whole expression tends to 1. ■

VI. EXPERIMENTAL EVALUATION

In this section we perform two experiments to illustrate the three results: Proposition 2, Proposition 4 and Proposition 5. The goal of these experiments is to compare the analytical

results with observations from a simulation process. Since there is no closed form result for ϵ ; this entity has to be estimated. We estimate ϵ in the same way as in the second test of Section III-A, and by averaging over all signals. To find the performance of the different voting strategies, we count how many times ‘ h ’ hits and ‘ m ’ misses correspond to a correct support-set index estimate and divide this number with the number of times ‘ h ’ hits and ‘ m ’ misses occurs in total. Thus the procedure is as follows:

- 1) Estimate $\hat{\epsilon}$.
- 2) For each $\hat{\epsilon}$, count the actual accuracy of the voting procedure and put a mark at this point.
- 3) Compare to the theoretical expression in the respective equation.

In Fig. 5, ϵ is plotted vs the probability of detection for the results of the common support-set model. A total number of 10^6 random simulations are performed, using parameters $N = 1000$, $T = 20$ and GP algorithm subspace pursuit (SP). To find different ϵ , M and SMNR are varied. In the case where $h = 2, m = 1$, the ϵ from left to right are found by: $M = 96, 85, 76, 64, 50, 41, 34, 28$ with corresponding SMNR = 20, 20, 20, 10, 10, 10, 10, 0 and for the case where $h = 3, m = 7$, $M = 101, 96, 92, 88, 50, 41, 34, 28$ with corresponding SMNR = 20, 20, 20, 20, 10, 10, 10, 0. Observe that largest possible $\epsilon_{\max} = \frac{N-T}{N}$ (marked with a small black dot). The equations used for the analytical results are found in (13). When we compare the simulations to the result predicted by analysis, we notice an almost perfect match. We argue that the slight mismatch for some points is due to noise and will average out using a larger simulation ensemble. For example, it is a rare event that $h = 3, m = 7$ occurs when the algorithms are very good (i.e., the simulation point at (0.05, 0.05)).

In Fig. 6, ϵ is plotted vs the probability of detection for the mixed support-set model. A total number of 10^6 random simulations are performed, using parameters $N = 1000$ and GP algorithm subspace pursuit (SP). To find different ϵ , M and SMNR are varied. Here we used the same data-points for all curves; the ϵ from left to right: $M = 96, 90, 85, 76, 64, 50, 41, 34, 28$ with corresponding SMNR = 20, 20, 20, 20, 10, 10, 10, 10, 0. We observe that also here, the simulation points match closely to the predicted values. The equations used for the analytical results are found in (15) and

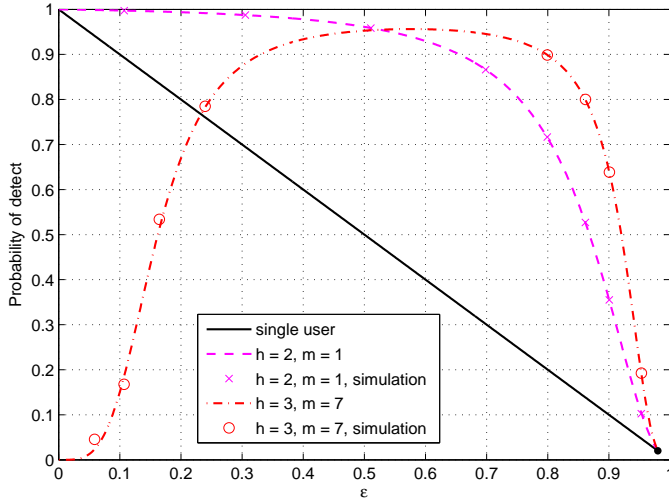


Fig. 7. Analytical and simulation results for the voting performance based on the common support-set model.

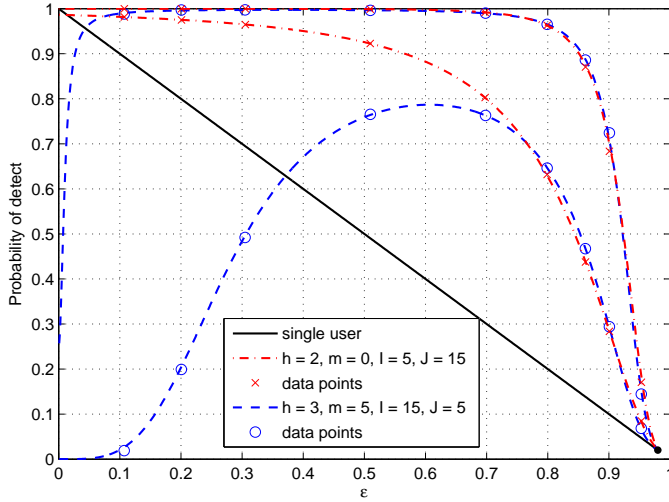


Fig. 8. Analytical and simulation results for the consensus performance based on the mixed support-set model.

(16).

VII. CONCLUSION

In this paper, we have analyzed democratic based voting strategies for support-sets estimation using greedy algorithms. We have characterized the input/output relation of any typical GP algorithm based on four relations. Using these relations we shown the merit of voting for two particular examples: the majority algorithm and the consensus algorithm, both which has been presented in the literature earlier. With several experiments, we validated both the input/output relation and the results derived; in all cases the experiments closely matched the theoretical prediction.

APPENDIX A PROOFS

Here, we present proofs for the propositions provided in the paper. First, we introduce some lemmas used in the proofs.

Then, in appendix A-A we present the proofs for proposition 2 and proposition 3; in appendix A-B we present the proofs for proposition 4, proposition 5 and remark 1.

Lemma 1 (Equi-probability of Subsets): For any $\mathcal{A} \subseteq \mathcal{T}_p$ and for any $\mathcal{B} \subseteq \mathcal{T}_p^c$, the following holds

$$\mathbb{P}(i \in \mathcal{A} | i \in \hat{\mathcal{T}}_p) = \frac{|\mathcal{A}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p), \quad (20)$$

$$\mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{A}) = \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p), \quad (21)$$

$$\mathbb{P}(i \in \mathcal{B} | i \in \hat{\mathcal{T}}_p) = \frac{|\mathcal{B}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c), \quad (22)$$

$$\mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{B}) = \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c). \quad (23)$$

Proof for (20):

$$\begin{aligned} \mathbb{P}(i \in \mathcal{A} | i \in \hat{\mathcal{T}}_p) &\stackrel{(a)}{=} \frac{|\mathcal{A}|}{T} \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p) \\ &= \frac{|\mathcal{A}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p) \frac{\mathbb{P}(i \in \mathcal{T}_p)}{\mathbb{P}(i \in \hat{\mathcal{T}}_p)} \\ &= \frac{|\mathcal{A}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p) \end{aligned}$$

Here, (a) follows directly from Definition 1. ■

Proof for (21):

$$\begin{aligned} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{A}) &= \mathbb{P}(i \in \mathcal{A} | i \in \hat{\mathcal{T}}_p) \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p)}{\mathbb{P}(i \in \mathcal{A})} \\ &\stackrel{(a)}{=} \frac{|\mathcal{A}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p) \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p)}{\mathbb{P}(i \in \mathcal{A})} \\ &= \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p), \end{aligned}$$

where we in (a) applied (20). ■

Proof for (22): This proof is similar to the proof for (20),

$$\begin{aligned} \mathbb{P}(i \in \mathcal{B} | i \in \hat{\mathcal{T}}_p) &= \frac{|\mathcal{B}|}{|\mathcal{T}_p^c|} \mathbb{P}(i \in \mathcal{T}_p^c | i \in \hat{\mathcal{T}}_p) \\ &= \frac{|\mathcal{B}|}{|\mathcal{T}_p^c|} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c) \frac{|\mathcal{T}_p^c|}{|\hat{\mathcal{T}}_p|} \\ &= \frac{|\mathcal{B}|}{T} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c). \end{aligned}$$

Proof for (23): This proof is similar to the proof for (21) and follows directly by applying (22). ■

Lemma 2 (Independence of Joint Probability): The local results from different sensor nodes are independent over certain regions. Assume there are a total of $h + m$ nodes in the system and that we denote different nodes by sub-indices $p_k \neq q_l \forall k, l$ and $p \neq p_k, p \neq q_l \forall k, l$. Then, for $\mathcal{A} \subseteq \mathcal{J}$, $\mathcal{B} \subseteq \mathcal{I}_{p_h}$, $\mathcal{C} \subseteq \mathcal{I}_{q_m}$, and $\mathcal{D} \subseteq (\mathcal{J} \cup \bigcup_{l=1}^h \mathcal{I}_{p_l} \cup \bigcup_{l=1}^m \mathcal{I}_{q_l})^c$ the following relations hold:

$$\begin{aligned} \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{A}) &\quad (24) \\ &= \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p)^h \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p)^m, \end{aligned}$$

$$\begin{aligned} \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{B}) &\quad (25) \\ &= \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c)^{h-1} \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p^c)^m, \end{aligned}$$

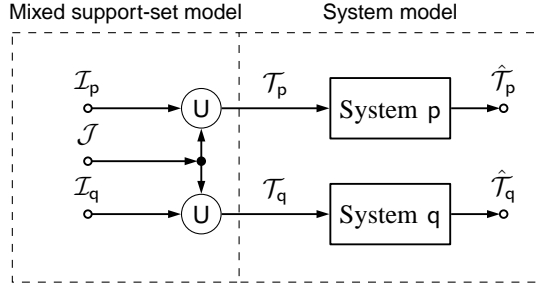


Fig. 9. For two nodes, this figure illustrates the first order Markov property of the outputs \hat{T}_p and \hat{T}_q .

$$\mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{C}) \quad (26)$$

$$= \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p^c)^h \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c)^{m-1},$$

$$\mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{D}) \quad (27)$$

$$= \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p^c)^h \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c)^m.$$

Proof of (24):

$$\begin{aligned} & \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{A}) \\ & \stackrel{(a)}{=} \prod_{l=1}^h \mathbb{P}(i \in \hat{T}_{p_l} | i \in \mathcal{A}) \prod_{l=1}^m \mathbb{P}(i \in \hat{T}_{q_l}^c | i \in \mathcal{A}), \\ & \stackrel{(b)}{=} \prod_{l=1}^h \mathbb{P}(i \in \hat{T}_{p_l} | i \in \mathcal{T}_{p_l}) \prod_{l=1}^m \mathbb{P}(i \in \hat{T}_{q_l}^c | i \in \mathcal{T}_{q_l}), \\ & \stackrel{(c)}{=} \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p)^h \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p)^m, \end{aligned}$$

In (a) we applied the chain rule on all intersections and applying the Markov property, which is illustrated in Fig. 9. In (b) we have used Lemma 1 and in (c) we have used that all probabilities are equal. ■

Proof of (25):

$$\mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{B}) \quad (28)$$

$$= \mathbb{P}(i \in \hat{T}_{p_h}, i \in \bigcap_{l=1}^{h-1} \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{B})$$

$$= \mathbb{P}(i \in \hat{T}_{p_h} | i \in \mathcal{B}) \prod_{l=1}^{h-1} \mathbb{P}(i \in \hat{T}_{p_l} | i \in \mathcal{B})$$

$$\cdot \prod_{l=1}^m \mathbb{P}(i \in \bigcap \hat{T}_{q_l}^c | i \in \mathcal{B})$$

$$\stackrel{(c)}{=} \mathbb{P}(i \in \hat{T}_{p_h} | i \in \mathcal{T}_{p_h}) \prod_{l=1}^{h-1} \mathbb{P}(i \in \hat{T}_{p_l} | i \in \mathcal{T}_{p_l}^c)$$

$$\cdot \prod_{l=1}^m \mathbb{P}(i \in \bigcap \hat{T}_{q_l}^c | i \in \mathcal{T}_{q_l}^c)$$

$$\stackrel{(d)}{=} \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p^c)^{h-1} \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c)^m.$$

In (a), Lemma 1 is used, and in (b) we used that all probabilities are equal. ■

Proof of (26): The proofs is similar to the proof for (25). ■

Proof of (27): The proofs is similar to the proof for (24). ■

Lemma 3 (Joint Probability): Assume there are $h+m$ nodes in the system, that $p_k \neq q_l \forall k, l$ are different nodes. Then, the following holds:

$$\begin{aligned} & \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c) \\ & = (1 - \epsilon)^h \epsilon^m \frac{J}{N} \\ & \quad + h(1 - \epsilon) \left(\frac{T}{N - T} \epsilon \right)^{h-1} \left(1 - \frac{T}{N - T} \epsilon \right)^m \frac{I}{N} \\ & \quad + m \epsilon \left(\frac{T}{N - T} \epsilon \right)^h \left(1 - \frac{T}{N - T} \epsilon \right)^{m-1} \frac{I}{N} \\ & \quad + \left(\frac{T}{N - T} \epsilon \right)^h \left(1 - \frac{T}{N - T} \epsilon \right)^m \frac{N - J - (m + h)I}{N} \end{aligned}$$

Proof: For this proof, we first introduce a notational simplification, define $\mathcal{U} = \mathcal{J} \cup \mathcal{I}_{p_1} \cup \dots \cup \mathcal{I}_{p_h} \cup \mathcal{I}_{q_1} \cup \dots \cup \mathcal{I}_{q_m}$. Observe that the sub-sets in \mathcal{U} are non-overlapping. Then,

$$\begin{aligned} & \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c) \\ & \stackrel{(a)}{=} \sum_{\substack{\mathcal{J}, \\ \mathcal{A} = \mathcal{I}_{p_1}, \mathcal{I}_{p_2}, \dots, \mathcal{I}_{p_h}, \\ \mathcal{U}^c = \mathcal{I}_{q_1}, \mathcal{I}_{q_2}, \dots, \mathcal{I}_{q_m}, \\ \mathcal{U}^c}} \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{A}) \mathbb{P}(i \in \mathcal{A}) \\ & \stackrel{(b)}{=} \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{J}) \mathbb{P}(i \in \mathcal{J}) \\ & \quad + h \mathbb{P}(i \in \hat{T}_p, i \in \bigcap_{l=1}^{h-1} \hat{T}_{q_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{I}_p) \mathbb{P}(i \in \mathcal{I}_p) \\ & \quad + m \mathbb{P}(i \in \hat{T}_p^c, i \in \bigcap_{l=1}^h \hat{T}_{q_l}, i \in \bigcap_{l=1}^{m-1} \hat{T}_{q_l}^c | i \in \mathcal{I}_p) \mathbb{P}(i \in \mathcal{I}_p) \\ & \quad + \mathbb{P}(i \in \bigcap_{l=1}^h \hat{T}_{p_l}, i \in \bigcap_{l=1}^m \hat{T}_{q_l}^c | i \in \mathcal{U}^c) \mathbb{P}(i \in \mathcal{U}^c) \\ & \stackrel{(c)}{=} \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p)^h \mathbb{P}(i \in \hat{T}_q^c | i \in \mathcal{T}_q)^m \mathbb{P}(i \in \mathcal{J}) \\ & \quad + h \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{T}_q | i \in \mathcal{T}_q)^{h-1} \\ & \quad \cdot \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c)^m \mathbb{P}(i \in \mathcal{I}_p) \\ & \quad + m \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c) \mathbb{P}(i \in \hat{T}_q | i \in \mathcal{T}_q)^h \\ & \quad \cdot \mathbb{P}(i \in \hat{T}_p^c | i \in \mathcal{T}_p^c)^{m-1} \mathbb{P}(i \in \mathcal{I}_p) \\ & \quad + \mathbb{P}(i \in \hat{T}_p | i \in \mathcal{T}_p^c)^h \mathbb{P}(i \in \hat{T}_q^c | i \in \mathcal{T}_q^c)^m \mathbb{P}(i \in \mathcal{U}^c) \\ & \stackrel{(d)}{=} (1 - \epsilon)^h \epsilon^m \frac{J}{N} \\ & \quad + h(1 - \epsilon) \left(\frac{T}{N - T} \epsilon \right)^{h-1} \left(1 - \frac{T}{N - T} \epsilon \right)^m \frac{I}{N} \\ & \quad + m \epsilon \left(\frac{T}{N - T} \epsilon \right)^h \left(1 - \frac{T}{N - T} \epsilon \right)^{m-1} \frac{I}{N} \end{aligned}$$

$$+ \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{N-J-(m+h)I}{N}.$$

In (a), the probability is marginalized over all individual and joint support-sets, and over \mathcal{U} . In (b), we extend the sum. In (c) we apply Lemma 2. Lastly, in (d) we plug the values from Definition 1. ■

A. Proofs of the Results for majority

Here, we prove Proposition 2 and Proposition 3, which are stated based on the common support-set model. Recall that in the common support-set model, $\mathcal{J} = \mathcal{T}$ and $J = T$.

Proof of Proposition 2:

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ &= \frac{\mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{J}) \mathbb{P}(i \in \mathcal{J})}{\mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c)} \\ &= \frac{\mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{J}) \mathbb{P}(i \in \mathcal{J})}{\sum_{\mathcal{A}=\mathcal{J}, \mathcal{J}^c} \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{A}) \mathbb{P}(i \in \mathcal{A})} \\ &\stackrel{(a)}{=} \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p)^h \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p)^m \mathbb{P}(i \in \mathcal{J})}{\sum_{\mathcal{A}=\mathcal{J}, \mathcal{J}^c} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{A})^h \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{A})^m \mathbb{P}(i \in \mathcal{A})} \\ &\stackrel{(b)}{=} \frac{(1-\epsilon)^h \epsilon^m \frac{J}{\mathcal{N}}}{(1-\epsilon)^h \epsilon^m \frac{J}{\mathcal{N}} + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m \frac{N-J}{\mathcal{N}}} \quad (29) \end{aligned}$$

In (a) Lemma 2 is applied, and lastly for (b), Definition 1 is used. ■

Proof of Proposition 3: This proposition states that he following inequality holds

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ & \geq \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^{h-1} \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{q_l}^c) \quad (30) \end{aligned}$$

Using Proposition 2, the LHS of (30) is:

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ &= \frac{(1-\epsilon)^h \epsilon^m J}{(1-\epsilon)^h \epsilon^m J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J)} \quad (31) \end{aligned}$$

Similarly, the RHS of (30) is

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \bigcap_{l=1}^{h-1} \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{q_l}^c) \\ &= \frac{(1-\epsilon)^{h-1} \epsilon^{m+1} J}{(1-\epsilon)^{h-1} \epsilon^{m+1} J + \left(\frac{T}{N-T}\epsilon\right)^{h-1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} (N-J)} \quad (32) \end{aligned}$$

By plugging the (31) and (32) into the inequality (30) we get:

$$\begin{aligned} & \frac{(1-\epsilon)^h \epsilon^m J}{(1-\epsilon)^h \epsilon^m J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J)} \\ & \geq \end{aligned}$$

$$\frac{(1-\epsilon)^{h-1} \epsilon^{m+1} J}{(1-\epsilon)^{h-1} \epsilon^{m+1} J + \left(\frac{T}{N-T}\epsilon\right)^{h-1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} (N-J)}$$

Multiplying the denominators gives

$$\begin{aligned} & (1-\epsilon) \left((1-\epsilon)^{h-1} \epsilon^{m+1} J \right. \\ & \quad \left. + \left(\frac{T}{N-T}\epsilon\right)^{h-1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} (N-J) \right) \\ & \geq \\ & \epsilon \left((1-\epsilon)^h \epsilon^m J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J) \right), \end{aligned}$$

which by simplifying gives

$$\begin{aligned} & \frac{(1-\epsilon)^h \epsilon^{m+1} J}{(1-\epsilon)^h \epsilon^{m+1} J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J)} \\ & \geq \\ & \frac{(1-\epsilon)^h \epsilon^{m+1} J}{(1-\epsilon)^h \epsilon^{m+1} J + \left(\frac{T}{N-T}\epsilon\right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J)}. \end{aligned}$$

Further simplifications give

$$\begin{aligned} & (1-\epsilon) \left(\frac{T}{N-T}\epsilon \right)^{h-1} \left(1 - \frac{T}{N-T}\epsilon\right)^{m+1} (N-J) \\ & \geq \\ & \epsilon \left(\frac{T}{N-T}\epsilon \right)^h \left(1 - \frac{T}{N-T}\epsilon\right)^m (N-J). \end{aligned}$$

Thus, we arrive at

$$\left(1 - \frac{T}{N-T}\epsilon\right)(1-\epsilon) \geq \frac{T}{N-T}\epsilon^2,$$

which can be simplified to

$$1 - \frac{N}{N-T}\epsilon + \frac{T}{N-T}\epsilon^2 \geq \frac{T}{N-T}\epsilon^2.$$

This is in turn equivalent to

$$\frac{N}{N-T}\epsilon \leq 1,$$

where the expression reaches its maximum at $\epsilon_{\max} = \frac{N-T}{N}$

$$\frac{N}{N-T}\epsilon \leq \frac{N}{N-T} \frac{N-T}{N} = 1.$$

Thus, we conclude that the sought inequality (14) holds true. ■

B. Proof of the Results for consensus

We now prove Proposition 4, Proposition 5 and Remark 1, which are based on the mixed support-set model.

Proof of Proposition 4: We first notice that the problem can be split into two parts,

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \mathbb{P}(i \in \mathcal{J} | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \quad (33) \end{aligned}$$

$$+ \mathbb{P}(i \in \mathcal{I}_p | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c), \quad (34)$$

where we consider each part separately.

- First we study (33)

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{J}) \mathbb{P}(i \in \mathcal{J})}{\mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c)}. \end{aligned} \quad (35)$$

We now consider each probability in (35) separately, beginning with

$$\begin{aligned} & \mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{J}) \\ &= \mathbb{P}(i \in \bigcap_{l=1}^{h+1} \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c | i \in \mathcal{J}) \\ &\stackrel{(a)}{=} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p)^{h+1} \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p)^m \\ &\stackrel{(b)}{=} (1 - \epsilon)^{h+1} \epsilon^m. \end{aligned} \quad (36)$$

Here, Lemma 2 was used in (a) and Definition 1 was used in (b). We have from the uniformity of the support-sets that

$$\mathbb{P}(i \in \mathcal{J}) = \frac{J}{N}. \quad (37)$$

Finally we have

$$\begin{aligned} & \mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \mathbb{P}(i \in \bigcap_{l=1}^{h+1} \hat{\mathcal{T}}_{p_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{q_l}^c) \\ &\stackrel{(a)}{=} (1 - \epsilon)^{h+1} \epsilon^m \frac{J}{N} \\ &\quad + (h+1)(1 - \epsilon) \left(\frac{T}{N-T} \epsilon \right)^h \left(1 - \frac{T}{N-T} \epsilon \right)^m \frac{I}{N} \\ &\quad + m \epsilon \left(\frac{T}{N-T} \epsilon \right)^{h+1} \left(1 - \frac{T}{N-T} \epsilon \right)^{m-1} \frac{I}{N} \\ &\quad + \left(\frac{T}{N-T} \epsilon \right)^{h+1} \left(1 - \frac{T}{N-T} \epsilon \right)^m \frac{N-J-(m+h+1)I}{N}, \end{aligned} \quad (38)$$

where (a) is achieved by Lemma 3.

- We now study (34)

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{I}_p | i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{I}_p) \mathbb{P}(i \in \mathcal{I}_p)}{\mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c)}. \end{aligned} \quad (39)$$

We now consider each probability in (39) separately, beginning with

$$\mathbb{P}(i \in \hat{\mathcal{T}}_p, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{I}_p)$$

$$\begin{aligned} & \stackrel{(a)}{=} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p^c)^h \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p^c)^m \\ & \stackrel{(b)}{=} (1 - \epsilon) \left(\frac{T}{N-T} \epsilon \right)^h \left(1 - \left(\frac{T}{N-T} \epsilon \right) \right)^m, \end{aligned}$$

where we just as for (36), used Lemma 2 for (a) and Definition 1 for (b). We have from the uniformity of support-sets that,

$$\mathbb{P}(i \in \mathcal{I}_p) = \frac{I}{N}. \quad (40)$$

Finally we notice for the third probability that the denominator is identical to (38). Now plugging the parts together gives (15). ■

Proof of Proposition 5: This proof is similar to the proof of Proposition 4. First, split the problem into two parts,

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \mathbb{P}(i \in \mathcal{J} | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \end{aligned} \quad (41)$$

$$+ \mathbb{P}(i \in \mathcal{I}_p | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c), \quad (42)$$

and we study each part separately.

- First study (41)

$$\begin{aligned} & \mathbb{P}(i \in \mathcal{J} | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\ &= \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{J}) \mathbb{P}(i \in \mathcal{J})}{\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c)}. \end{aligned} \quad (43)$$

This was achieved using Bayes' rule. We now consider each probability in (43) separately, beginning with

$$\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{J}) \quad (44)$$

$$= \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{J}) \quad (45)$$

$$\stackrel{(a)}{=} \mathbb{P}(i \in \hat{\mathcal{T}}_p | i \in \mathcal{T}_p)^h \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p)^{m+1} \quad (46)$$

$$\stackrel{(b)}{=} (1 - \epsilon)^h \epsilon^{m+1} \quad (47)$$

Here, Lemma 2 was used in (a) and Definition 1 was used in (b). We have from the uniformity of the support-sets that

$$\mathbb{P}(i \in \mathcal{J}) = \frac{J}{N}. \quad (48)$$

Finally we have

$$\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \quad (49)$$

$$= \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{r_l}^c) \quad (50)$$

$$\begin{aligned}
&\stackrel{(a)}{=} (1-\epsilon)^h \epsilon^{m+1} \frac{J}{N} \\
&\quad + h(1-\epsilon) \left(\frac{T}{N-T} \epsilon \right)^{h-1} \left(1 - \frac{T}{N-T} \epsilon \right)^{m+1} \frac{I}{N} \\
&\quad + (m+1) \epsilon \left(\frac{T}{N-T} \epsilon \right)^h \left(1 - \frac{T}{N-T} \epsilon \right)^m \frac{I}{N} \\
&\quad + \left(\frac{T}{N-T} \epsilon \right)^h \left(1 - \frac{T}{N-T} \epsilon \right)^{m+1} \frac{N-J-(m+h+1)I}{N},
\end{aligned} \tag{51}$$

where (a) is obtained by Lemma 3.

- For (42) we have that

$$\begin{aligned}
&\mathbb{P}(i \in \mathcal{I}_p | i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c) \\
&= \frac{\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{I}_p) \mathbb{P}(i \in \mathcal{I}_p)}{\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c)},
\end{aligned} \tag{52}$$

which is achieved with Bayes' rule. We now consider each probability in (52) separately, beginning with the first probability

$$\begin{aligned}
&\mathbb{P}(i \in \hat{\mathcal{T}}_p^c, i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^m \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{I}_p) \\
&= \mathbb{P}(i \in \bigcap_{l=1}^h \hat{\mathcal{T}}_{q_l}, i \in \bigcap_{l=1}^{m+1} \hat{\mathcal{T}}_{r_l}^c | i \in \mathcal{I}_{r_{m+1}}) \\
&\stackrel{(a)}{=} \mathbb{P}(i \in \hat{\mathcal{T}}_p^c | i \in \mathcal{T}_p) \mathbb{P}(i \in \hat{\mathcal{T}}_q | i \in \mathcal{T}_q^c)^h \mathbb{P}(\hat{\mathcal{T}}_r^c | i \in \mathcal{T}_r^c)^m \\
&\stackrel{(b)}{=} \epsilon \left(\frac{T}{N-T} \epsilon \right)^h \left(1 - \frac{T}{N-T} \epsilon \right)^m,
\end{aligned}$$

where we used Lemma 2 for (a) and Definition 1 for (b). We have from the uniformity of support-sets that,

$$\mathbb{P}(i \in \mathcal{I}_p) = \frac{I}{N}.$$

Finally we notice for the third probability that the denominator is identical to (51). Now plugging all the parts together gives (16). \blacksquare

Proof of Remark 1: We first study (17) and notice that any index in $i \in \hat{\mathcal{T}}_p$ fulfills one of the following: $i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)$, or $i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r^c)$, or $i \in (\hat{\mathcal{T}}_p^c \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)$. Thus, we will show the remark by proving each of the following inequalities:

$$\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)) \geq \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p), \tag{53}$$

$$\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r^c)) \geq \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p), \tag{54}$$

$$\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p^c \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)) \geq \mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p). \tag{55}$$

First, recall from Proposition 1 and (9) that

$$\mathbb{P}(i \in \mathcal{T}_p | i \in \hat{\mathcal{T}}_p) = 1 - \epsilon. \tag{56}$$

- We now consider (53). By plugging $N = 1000$, $T = 20$, $J = 15$, $I = 5$, $m = 0$ and $h = 2$ into Proposition 4, we obtain

$$\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)) \tag{57}$$

$$= \frac{49(\epsilon - 1)(7204\epsilon^2 - 14406\epsilon + 7203)}{352900\epsilon^3 - 1058988\epsilon^2 + 1058841\epsilon - 352947}.$$

We multiply the denominator of (57) to (56) get the following inequality

$$\begin{aligned}
&49(\epsilon - 1)(7204\epsilon^2 - 14406\epsilon + 7203) \\
&\geq (1 - \epsilon)(352900\epsilon^3 - 1058988\epsilon^2 + 1058841\epsilon - 352947),
\end{aligned}$$

which equivalently can be simplified to

$$0 \geq (\epsilon(50\epsilon - 49)(7058\epsilon - 7203)(\epsilon - 1))/23529800. \tag{58}$$

The roots to the polynomial of (58) are: $\epsilon_1 = 0$, $\epsilon_2 = \frac{49}{50} = 0.98$, $\epsilon_3 = \frac{7203}{7058} = 1.0205\dots$ and $\epsilon_4 = 1$. Thus, the interesting region is $\epsilon \in [\epsilon_1, \epsilon_2]$, for which the inequality (58) holds.

- We now consider (54). By plugging $N = 1000$, $T = 20$, $J = 15$, $I = 5$, $m = 1$ and $h = 1$ into Proposition 4, we obtain

$$\begin{aligned}
&\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r^c)) \\
&= \frac{196\epsilon(1801\epsilon^2 - 3614\epsilon + 1813)}{\epsilon(352900\epsilon^2 - 701288\epsilon + 357749)}.
\end{aligned} \tag{59}$$

Observe that also here, $\epsilon = 0$ is undefined. We multiply the denominator of (59) to (56) and get the following inequality

$$\begin{aligned}
&196\epsilon(1801\epsilon^2 - 3614\epsilon + 1813) \\
&\geq \epsilon(1 - \epsilon)(352900\epsilon^2 - 701288\epsilon + 357749),
\end{aligned}$$

which can be simplified to

$$0 \geq -(\epsilon(50\epsilon - 49)(7058\epsilon - 49)(\epsilon - 1))/23529800. \tag{60}$$

The roots to the polynomial of (60) are: $\epsilon_1 = 0$, $\epsilon_2 = \frac{49}{50} = 0.98$, $\epsilon_3 = \frac{49}{7058} = 0.0069\dots$ and $\epsilon_4 = 1$. Thus, the interesting region is $\epsilon \in [\epsilon_3, \epsilon_4]$, for which the inequality (60) holds.

- We now consider (55). By plugging $N = 1000$, $T = 20$, $J = 15$, $I = 5$, $m = 0$ and $h = 2$ into Proposition 5, we obtain

$$\begin{aligned}
&\mathbb{P}(i \in \mathcal{T}_p | i \in (\hat{\mathcal{T}}_p^c \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r)) \\
&= \frac{49\epsilon(7204\epsilon^2 - 14406\epsilon + 7203)}{\epsilon(352900\epsilon^2 - 701288\epsilon + 357749)}.
\end{aligned} \tag{61}$$

Observe that in this expression, $\epsilon = 0$ is undefined. This is naturally true² and follows directly from Lemma 3. We multiply the denominator of (61) to (56) and get the following inequality

$$\begin{aligned}
&49\epsilon(7204\epsilon^2 - 14406\epsilon + 7203) \\
&\geq \epsilon(1 - \epsilon)(352900\epsilon^2 - 701288\epsilon + 357749),
\end{aligned}$$

which can be simplified to

$$0 \geq -(\epsilon(50\epsilon - 49)(7058\epsilon^2 - 7107\epsilon + 98))/23529800. \tag{62}$$

²If all algorithms are perfect, the cut $(\hat{\mathcal{T}}_p \cap \hat{\mathcal{T}}_q \cap \hat{\mathcal{T}}_r^c) = \emptyset$.

The roots to the polynomial of (62) are: $\epsilon_1 = 0$, $\epsilon_2 = \frac{49}{50} = 0.98$, $\epsilon_3 = \frac{7107}{2 \cdot 7058} + \sqrt{\left(\frac{7107}{2 \cdot 7058}\right)^2 - \frac{98}{7058}} = 0.9930\dots$ and $\epsilon_4 = \frac{7107}{2 \cdot 7058} - \sqrt{\left(\frac{7107}{2 \cdot 7058}\right)^2 - \frac{98}{7058}} = 0.0140\dots$. Thus, the interesting region is $\epsilon \in [\epsilon_4, \epsilon_3]$, for which the inequality (62) holds.

From the above calculations, we find the interesting region is the region that lies between $\epsilon_4 \geq 0.0140$ for (62) and $\epsilon_{\max} = \frac{N-T}{N} = 0.98$. Since all the inequalities (58), (60) or (62) hold true in this region (directly verified by plugging in any $0.0140 \leq \epsilon \leq 0.98$), we conclude the proof. ■

REFERENCES

- [1] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, Apr. 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, Aug. 2006.
- [3] J. Mota, J. Xavier, P. Aguiar, and M. Puschel, "Distributed basis pursuit," *IEEE Transactions on Signal Processing*, vol. 60, pp. 1942–1956, Apr. 2012.
- [4] J. Bazerque and G. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1847–1862, Mar. 2010.
- [5] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, pp. 2346–2356, Jun. 2008.
- [6] D. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 18914–18919, Oct. 2009.
- [7] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, Dec. 1993.
- [8] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, pp. 4655–4666, Dec. 2007.
- [9] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, pp. 301–321, Apr. 2009.
- [10] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, pp. 2230–2249, May 2009.
- [11] D. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, pp. 1094–1121, Feb. 2012.
- [12] S. Chatterjee, D. Sundman, M. Vehkaperä, and M. Skoglund, "Projection-based and look-ahead strategies for atom selection," *IEEE Transactions on Signal Processing*, vol. 60, pp. 634–647, Feb. 2012.
- [13] D. Sundman, S. Chatterjee, and M. Skoglund, "FROGS: A serial reversible greedy search algorithm," in *IEEE Swedish Communication Technologies Workshop (Swe-CTW)*, Lund, Sweden, Mar. 2012, pp. 40–45.
- [14] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, pp. 310–316, Mar. 2010.
- [15] D. Sundman, S. Chatterjee, and M. Skoglund, "Look ahead parallel pursuit," in *IEEE Swedish Communication Technologies Workshop (Swe-CTW)*, Stockholm, Sweden, Oct. 2011, pp. 114–117.
- [16] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, vol. 346, pp. 589–592, May 2008.
- [17] J. Tropp, A. Gilbert, and M. Strauss, "Algorithms for simultaneous sparse approximation. part i: Greedy pursuit," *Signal Processing*, vol. 86, pp. 572–588, Mar. 2006.
- [18] A. Rakotomamonjy, "Surveying and comparing simultaneous sparse approximation (or group-lasso) algorithms," *Signal Processing*, vol. 91, pp. 1505–1526, Jul. 2011.
- [19] D. Leviatan and V. Temlyakov, "Simultaneous approximation by greedy algorithms," *Advances in Computational Mathematics*, vol. 25, pp. 73–90, Jul. 2006.
- [20] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2477–2488, Jul. 2005.
- [21] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4634–4643, Dec. 2006.
- [22] D. Sundman, S. Chatterjee, and M. Skoglund, "Greedy pursuits for compressed sensing of jointly sparse signals," in *EURASIP European Signal Processing Conference (EUSIPCO)*, Barcelona, Spain, Aug. 2011, pp. 368–372.
- [23] J. Bazerque and G. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1847–1862, Mar. 2010.
- [24] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," vol. 5, pp. 37–48, Feb. 2011.
- [25] Q. Ling and T. Zhi, "Decentralized support detection of multiple measurement vectors with joint sparsity," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 2996–2999.
- [26] D. Sundman, D. Zachariah, and S. Chatterjee, "Distributed predictive subspace pursuit," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 4633–4637.
- [27] D. Sundman, S. Chatterjee, and M. Skoglund, "A greedy pursuit algorithm for distributed compressed sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 2729–2732.
- [28] —, "Distributed greedy pursuit algorithms," *Signal Processing*, 2014, accepted.
- [29] —, "DIPP: A distributed greedy algorithm based on democratic voting principles," *CoRR*, vol. abs/1403.6974, May 2014.
- [30] —, "Methods for distributed compressed sensing," *MDPI Journal of Sensor and Actuator Networks*, vol. 3, pp. 1–25, Dec. 2013. [Online]. Available: <http://www.mdpi.com/2224-2708/3/1/1>
- [31] P. Young, "Optimal voting rules," *The Journal of Economic Perspectives*, vol. 9, pp. 51–64, Dec. 1995.
- [32] J. Ledyard and T. Palfrey, "The approximation of efficient public good mechanisms by simple voting schemes," *Journal of Public Economics*, vol. 83, pp. 153–171, Feb. 2002.
- [33] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, Nov. 2005, pp. 1537–1541.